NAME (from your UF ID): ______________________________ UF ID#: ________________
(Please **PRINT**)

---------------------------- CEN 4072/6070 Software Testing & Verification --------------------------

Exam 1 -- Spring 2017

You have 50 minutes to work on this exam. It is a "closed-book/closed-notes" test. Pay attention to point values, since you may not have time to work all 21 problems.

PRINT your name and UF ID# above NOW and sign the pledge at the bottom of this page, if appropriate, when you are finished.

PLEASE PRINT **– do NOT write** ***cursively*** **–** ANSWERS IN THE SPACE PROVIDED ONLY – **NOT IN THE MARGINS** – PREFERABLY USING A BALLPOINT PEN TO INCREASE LEGIBILITY. Good luck!

On my honor, I have neither given nor received unauthorized aid on this exam and I pledge not to divulge information regarding its contents to those who have not yet taken it.

____
SIGNATURE

1. (4 pts.) Two testing-related terms that directly relate to the discussion of alternative definitions of testing are “validation testing” and “defect testing”. Describe what these terms refer to, making the relevant distinction between them clear.

2. (3 pts.) In arguing that exhaustive input testing is generally impossible, we illustrated the problem using an example program in class. Which of the following was the example? (Circle ONE only.)

    a. a program module with 2 input parameters, each with a word size of 32 bits

    b. a word processing system

    c. a compiler

    d. a program that would determine if an ordered pair of numeric inputs are in ascending order, descending order, or equal

    e. an airline reservation system

    f. (none of the above)

3. (16 pts.) In “The Psychology and Economics of Program Testing,” Myers identifies a number of “vital testing principles” related to important considerations in software testing. Some of these are:

    A. A necessary part of a test case is the definition of the expected output or result.

    B. A programmer should avoid attempting to test his or her own program.

    C. A programming organization should not test its own programs.

    D. The results of each test case must be thoroughly inspected.

    E. Test cases must be written for invalid and unexpected as well as valid and expected, input conditions.

    F. Examining a program to see if it does not do what it is supposed to do is only half of the battle. The other half is seeing whether the program does what it is not supposed to do.

    G. Avoid throw-away test cases unless the program is truly a throw-away program.

    H. Do not plan a testing effort under the tacit assumption that no errors will be found.

    I. Testing is the process of executing a program with the intent of finding errors.

    J. A good test case is one that has a high probability of detecting an as yet undiscovered error.

    K. The probability of the existence of more errors in a section of a program is proportional to the number of errors already found in that section.

3. (cont'd)

For each description below, identify the single principle from among those listed above that best reflects what Myers is describing. (Enter the letter of the principle being described to the left of the description.)

_____ *This problem is seen most often in the use of interactive systems to test programs. A common practice is to sit at a terminal, invent test cases on the fly, and then send these test cases through the program...*

_____ *…the phenomenon has been observed in IBM operating systems. In one of these, 47% of the (errors found by users) were associated with only 4% of the modules within the system.*

_____ *…a project manager is largely measured on the ability to produce a program by a given date and for a certain cost. One reason for this is that it is easy to measure time and cost objectives, but it is extremely difficult to quantify the reliability of a program. Therefore it is difficult for a programming organization to be objective in testing its own program, because the testing process, if approached with the proper definition, may be viewed as decreasing the probability of meeting the schedule and cost objectives.*

_____ (this principle is related to the natural tendency that) *has frequently appeared in the testing of the triangle program in Chapter 1. Few people, for instance, feed the program the numbers 1,2,5 to make sure that the program does not erroneously interpret this as a scalene triangle.*

_____ *Quoting the logician Copi: A problem may be characterized as a fact or group of facts for which we have no acceptable explanation, which seem unusual, or which fail to fit in with our expectations or preconceptions. It should be obvious that some prior beliefs are required if anything is to appear problematic. If there are no expectations, there can be no surprises.*

_____ *This is probably the most obvious principle, but again it is something that is often overlooked. In experiments performed by the author, many subjects failed to detect certain errors, even when symptoms of those errors were clearly observable...*

_____ *This principle follows from earlier discussion in the chapter, principally the discussion that implied that testing is a destructive process. ...As many home-owners know, removing wall-paper (a destructive process) is not easy, but it is almost unbearably depressing if you, rather than someone else, originally installed it. ...In addition...there is a second significant problem: The program may contain errors due to the programmer's misunderstanding of the problem statement or specification...*

_____ *…if one were instructed to solve the crossword puzzle in the Sunday New York Times in 15 minutes, we would observe little, if any, progress after 10 minutes because the person would be resigned to the fact that the task seems impossible. If we asked for a solution in four hours, however, we would probably observe more progress in the initial 10 minutes.*

4. (12 pts.) For each of “The six essentials of software testing,” described by Edward Kit, circle the missing word(s) among the choices provided that he actually used. (Circle ONE only.)

a. Essential 1: **The quality of the ___________ determines the success of the test effort.**

i. testers
ii. test plan
iii. test tools
iv. test cases
v. test process
vi. test environment

b. Essential 2: **Prevent defect migration by using _________ techniques**.

i. information hiding
ii. incremental integration testing
iii. regression testing
iv. effective change control
v. enhanced interrogation
vi. early life-cycle testing

c. Essential 3: **The time for __________ is now.**

i. reviews & inspections
ii. continuous process improvement
iii. formal methods
iv. tester training
v. software testing tools
vi. civil discourse

d. Essential 4: **A real person must take responsibility for __________** .

i. test planning
ii. software quality
iii. maintaining test cases
iv. improving the testing process
v. the QA function
vi. good snacks

e. Essential 5: **Testing is a professional discipline requiring __________** .

i. trained, skilled people
ii. disciplined self-deprivation
iii. a good sense of humor
iv. good intuition
v. access to drugs
vi. adequate compensation

f. Essential 6: **Cultivate a positive team attitude of __________** .

i. working together
ii. mutual respect
iii. disciplined invincibility
iv. quality improvement
v. creative destruction
vi. cultural relativism

5. (12 pts.) Match each description below to the **SINGLE MOST APPROPRIATE TERM** related to testing object-oriented software among the following. (Note: terms may apply to none, one, or more than one description.)

A. encapsulation
B. inheritance
C. unit level O-O testing
D. object classes
E. initial inter-object class testing
F. higher level O-O testing
G. polymorphism
H. observability interfaces
K. state machine models
L. methods
M. UML sequence or collaboration diagrams
N. inspection operations

___ Use/include relations allows the design of an incremental strategy

___ Object-oriented design principle that may necessitate the use of special instrumenttation to interpret the effects of methods

___ Often useful for describing object class states and transitions

___ Allows object classes to be *specialized* or *extended* from one or more other object classes

___ Feature is problematic when effects of operations are private and therefore “hidden” from the tester

___ Sometimes required by design rules to circumvent the problem of object state not being directly accessible, but results in some overhead

___ Allows variable types and method bindings to change dynamically, thus requiring testers to be aware of the bindings that may occur

___ Often focuses on object classes as opposed to individual methods, as testing methods in isolation is not always practical

___ Facilitates the application of partitioning and combinatorial test case design techniques on a state-by-state basis

___ An issue which arises as a result of this feature involves deciding which methods in a derived class need to be (re-)tested

___ Normally ***begins*** with the integration of object classes to form inter-object class functional entities (hint: best answer is NOT “E”)

___ Primary objective is to discover interface and blatant higher-level design errors

6. (2 pts.) As discussed in class, software reliability may be measured using *statistical testing* based on an operational profile. What is an “operational profile”?

7. (3 pts.) Recall the pseudocode program below discussed in class.

```
if (input is other than an ordered pair of numbers)
    then output(“invalid input”)
else
    if x<y then output(“ascending order”)
    else
         if x>y then output(“ascending order”)
         else
             output(“equal”)
```

Which one of the following statements best reflects the *primary* point made in connection with this program? (Circle ONE only.)

a. Correctness of the program can only be determined using statistical testing.

b. Correctness of the program cannot be determined in the absence of its requirements.

c. It illustrates a design for which the underlying assumption of partition testing would hold for the equivalence classes identified in class.

d. Correctness of the program can only be determined using BOTH black-box and white-box testing.

e. Attempting to count the "number of bugs/errors" in the program is probably not meaningful.

f. (none of the above)

8. (3 pts.) Recall the question posed in class regarding the "Fisherman's Dilemma:"

> *You have 3 days for fishing and 2 lakes to choose from. Day 1 at Lake X nets 8 fish. Day 2 at Lake Y nets 32 fish. Which lake do you return to for day 3?*

Which one of the following statements represents the *primary* point made about *testing* in connection with this issue? (Circle ONE only.)

a. Like fishing, testing is an extremely creative and challenging task.

b. In general, we can’t use testing to prove a program is error free for the same reason that we can’t use fishing to prove that a lake has no fish. Thus, in testing as in fishing, we need to weigh COST and RISK.

c. Both the cost of errors and the stench of dead fish tend to increase *exponentially* with time before discovery.

d. A programmer who is unable to avoid fishing for his own program errors should become Mr. Hyde... i.e., adopt a “tester’s mindset” that mitigates his ego-attachment to the program.

e. Testers should execute programs with the *intent* of finding errors for the same reason that fishermen should go fishing with the *intent* of catching fish.

f. (none of the above)

9. Microsoft’s Ken Johnston describes a new model of testing in his recorded Guest Lecture, “The Future of Testing is EaaS*y*.”

   a. (3 pts.) Briefly identify the model by name and describe the basic premise upon which it is based.

   b. (12 pts.) Briefly define and explain the relevance of each of the following terms/acronyms used by Ken Johnston to the model he describes.

      i. BUFT

      ii. LKGC

      iii. code churn

      iv. flighting

10. (3 pts.) In his paper, "Design and Code Inspections to reduce errors in program development," M.E. Fagan notes that in his experience with inspections, "evidently no manager has tried to 'kill the goose that lays the golden eggs'." Which one of the following describes the specific issue to which he was referring? (Circle ONE only.)

    a. Zealously ensuring that all issues, problems, and concerns discovered in an inspection have been resolved by the owner.

    b. The problem of including "bad eggs" (bad fixes or corrections) in a product due to the human tendency to consider a fix to be error-free itself.

    c. Pursuing questions raised during an inspection only to the point at which an error is recognized (i.e., without "killing the goose").

    d. Over-pressuring reviewers to seek high-occurrence, high cost error types (i.e., "golden eggs").

    e. The strong inclination of managers to change the adage that, "design is not complete until testing is completed."

    f. (none of the above)

11. Suppose a function that returns 2 Boolean values has 3 inputs, one of which is comprised of 4 disjoint attribute classes, while the others are comprised of 2 disjoint attribute classes each. There are no infeasible input combinations.

    a. (6 pts.) What is the minimum number of test cases needed to achieve *"Strong Equivalence Class Testing"*? What is the minimum number needed to achieve *"Weak Equivalence Class Testing"?*

    minimum # of cases needed for *strong equivalence class testing:* _____

    minimum # of cases needed for *weak equivalence class testing:* _____

    b. (3 pts.) Suppose a given implementation of this function has a total of 6 execution paths, 5 of which are feasible. If the function's input space is partitioned solely on the basis of *specified output,* what is the minimum number of test cases needed to ensure coverage of the partitions?

    minimum # of cases needed to cover the partitions: _____

12. (3 pts.) What potential problem with using techniques such as "Error Seeding" to estimate the number of errors remaining in a program during testing was illustrated in class by way of a pseudocode Quick Sort program? *How* did the program illustrate the problem? Be specific.

13. A variation on "Test Case Selection Strategy #3" that we considered involves the use of particular "culling rules".

a. (4 pts.) Describe the mechanics and specific combinatorial effect of applying this variation when encountering a 3rd-degree AND-node that must evaluate to False in the process of identifying Cause combinations resulting in a given Effect.

b. (3 pts.) Briefly explain the ***rationale*** for this particular culling approach.

14. (6 pts.) After just 2 days of testing a new product release, Janice happily announced to her manager that her team had already found 20 errors. The manager, who was surprised to hear that so many errors had been discovered so quickly, asked Janice why she was feeling happy in light of this depressing news. She then said, "Well, 16 of those 20 'errors' weren't real 'bugs' – they were, in fact, among 40 'errors' that I created and then 'seeded' into the system before testing started. So, I figure that there should only be just a few REAL bugs left to find now." Assuming Janice used the error seeding technique discussed in class, how many real bugs does she estimate are still left to find? (Show and explain your work.)

15. (3 pts.) During the discussion of testing tools, reference was made to the wonder and beauty of springtime in Paris…

Which ONE of the following testing tool types was this reference used to clandestinely introduce? (Circle ONE only.)

a. Boundary value, equivalence class, and special value generators

b. Statistical test generators (for reliability, availability, and performance testing)

c. Test Harness Generators

d. Data / Output Comparators

e. Keystroke Recorder and Playback Tools

f. Code Instrumentation tools

g. (none of the above)

16. (10 pts.) The following statements relate to Grady and Van Slack, et al., "Key Lessons in Achieving Widespread Inspection Use." Indicate whether each is true or false. To compensate for random guessing, your score in points will be 2 times the number of [correct minus incorrect] answers, or 0 – whichever is greater. Therefore, if you are not more than 50% sure of your answer, consider skipping the item.

a. The inspection process step that varies the most at HP is the "cause/prevention" step. true false

b. It was found that a particularly strong predictor of successful technology adoption at HP is the number of people trained in that technology. true false

c. While HP's culture leads to less formality than in many other companies, the inspection process emphasizes that this must NOT extend to the recording and tracking of meeting results. true false

d. HP created a measure called the Extent-Of-Adoption Metric to gauge company-wide progress in process maturity, depth of use, and breadth of use. true false

e. A major accomplishment that occurred during HP's "Widespread (Inspection) Belief and Adoption" stage (1989-1993) was the establishment of a single "standard" HP inspection process. true false

17. (4 pts.) In their paper, *Key Lessons in Achieving Widespread Inspection Use,* Grady and Van Slack motivate their story about the software inspection adoption process at Hewlett Packard with an anecdote that illustrates the enormous potential return on investment (ROI) that can be achieved with some types of inspections. Which one of the following was the subject of the anecdote given? (Circle ONE only.)

a. Before bringing a new hydroelectric power plant on-line for the first time, a $200 inspection of a generator bearing resulted in finding a previously undetected crack that would have resulted in the likely destruction of a $200,000 turbine.

b. Before buying a recently refurbished vintage sports car, one of the authors decided to invest in a $100 inspection of the car's crankshaft, despite assurances that it had recently been replaced. It turns out that the wrong part had been used during refurbishment, and that the engine would have likely been destroyed after just a few hours of use.

c. The authors describe the experience of one organization within HP that reported an average cost of $91 per defect found during "inspections" versus $25,000 per defect found after product delivery.

d. Before buying a used house, one of the authors and his wife decided to pay $25 for an inspection of the house's hot water based heating system. It turned out that the water pipes were leaking and the system had to be replaced at great expense – by the sellers.

e. The authors refer to an industry survey from 1985 that showed it is 100 times more expensive to correct errors discovered during "installation" than when discovered using inspections during "analysis."

f. (none of the above)

18. (10 pts.) The following statements relate to Sauer, et al., “The Effectiveness of Software Development Technical Reviews (SDTRs): A Behaviorally Motivated Program of Research.” Indicate whether each is true or false. To compensate for random guessing, your score in points will be 2 times the number of [correct minus incorrect] answers, or 0 – whichever is greater. Therefore, if you are not more than 50% sure of your answer, consider skipping the item.

a. The behavioral theory predicts that, where there is no plurality by which to decide whether an issue is a true defect, a group’s ability to make a correct discrimination is positively influence by the quality of its group processes. true false

b. The authors note that rather than ask whether or not SDTRs find more defects than no review at all, their concern is to ask *how much more effective they are than this most basic requirement.* true false

c. It was found that interacting groups generate a significant volume of new, creative problem solutions beyond those already generated in the individual phase of the task. true false

d. The most salient finding of the empirical research on which the theory is based is that group performance is dominated by the available task expertise. true false

e. The behavioral theory of group performance upon which the authors’ results are based stems from studies in which subjects are required to imagine themselves stranded in the desert with a limited number of implements available to them. true false

19. (4 pts.) In their paper, *The Effectiveness of Software Development Technical Reviews: A Behaviorally Motivated Program of Research,* Sauer, et al. apply the behavioral theory of group performance to explain the outcomes of software reviews. Which one of the following, according to the authors, is the “most salient (i.e., important) finding” of the empirical research in which the theory is based? (Circle ONE only.)

a. That decision schemes (plurality effects) dominate interacting group performance.

b. That the performance advantage of an interacting group over a nominal group is a function of the level of false positives discovered by individuals.

c. That synergy is the primary reason why groups perform better than individuals.

d. That group performance is a positive function of task training.

e. That group performance is dominated by the available task expertise.

f. (none of the above)

20. (20 pts.) In their chapter, *"Making Meetings Work for Everyone,"* Gause and Weinberg offer several observations and recommendations related to making meetings more productive. For each of the following, circle "is" if the observation or recommendation is included in the chapter, and "is not" otherwise. To compensate for random guessing, your score in points will be 2 times the number of [correct minus incorrect] answers, or 0 – whichever is greater. Therefore, if you are not more than 50% sure of your answer, consider skipping the item.

a. Limiting meetings to no more than 3 agenda items keeps them small, short, and relevant to all in attendance. is is not

b. People who inappropriately come to meetings are likely to start participating to "prove" they were right to be there in the first place. is is not

c. If a project seems to hold too many meetings, it may be a symptom of overstaffing. is is not

d. If the same people are to attend different meetings one right after another in the same room, have the participants drink a strong cup of coffee before starting the second meeting. is is not

e. Criticizing people for not attending meetings is likely to result in their becoming disruptive when they do attend in order to "prove" that it was a mistake to invite them in the first place. is is not

f. Meeting participants should agree in advance that anyone can call one-minute or five-minute time-outs at any point without explaining why. is is not

g. One way to gauge participant satisfaction is to simply ask each person, "Are you completely satisfied with each and every decisions made at this meeting?" is is not

h. Noticing what's wrong with your meetings will help diagnose what's wrong with the entire process, and indicate what to prescribe for the malady. is is not

i. One way to see if there is unfinished business at the end of a meeting is to ask each person, "Has every one of your ideas been handled to your satisfaction?" is is not

j. It is important to never allow agreed-upon rules to be changed during a meeting since this could be used as a political tool by one side against another. is is not

21. (18 pts.) Match each description below to the **SINGLE MOST APPROPRIATE TERM** among the following. (Note: terms may apply to none, one, or more than one description.)

A. Fault-based test
B. System test acceptance
C. Thread test
D. Alpha test
E. Performance test
F. Stress test
G. “Lights out” test
H. Regression test
I. Reliability test
J. Unit test
K. Benchmarking

L. “Soak” test
M. Recovery test
N. Usability test
O. Serviceability test
P. Beta test
R. Integration testing
S. “Smoke” test
T. Exhaustive test
U. Compatibility/conversion test
W. Causal analysis
X. Test-driven development

___ Focus is on typical requirements that systems exhibit “graceful” failures and non-abrupt performance degradation

___ Integrating program elements associated with a key program function

___ General practice of recording and comparing indices of performance, quality, cost, etc., to help identify “best practices”

___ Specialized testing in which HCI experts conduct experiments and utilize protocol analysis

___ Testing by end-users performed within the user environment prior to general release

___ (Limited) testing conducted to ensure that a system is “ready” (sufficiently stable) for the system-level test phase

___ Testing which takes place as sub-elements are combined to form higher-level elements

___ Re-testing to detect problems caused by the adverse effects of program change

___ Appropriate interpretations for “failure” and “time” are critical

___ Testing of the smallest programmer work assignments that can reasonably be planned and tracked

___ Automated, stand-alone testing not requiring human involvement

___ Testing a system version over a significant period of time to discover latent errors or performance problems

___ Introduced in support of agile methods; code is developed incrementally, along with one or more tests for each increment

___ May be automated by combining a keystroke recorder and playback tool with a data/output comparator

___ When all combinations of all possible input and state variable values are covered

___ Process aimed at identifying the origin of errors and approaches to eliminate future occurrences

___ Focus is on capabilities such as detecting exceptional conditions, switching over to standby systems, and maintaining audit trails

___ Typical coverage includes post-delivery change procedures (adaptive, perfective, and corrective scenarios), supporting documentation, and system diagnostic tools